Assignment of *CRSP9*, *ETF1* and *TMEM59* genes to porcine chromosomes

M.Y. Shao^{a, b} H. Wang^{a, b} Z.M. Zhu^b S.L. Yang^b K. Li^b

^aLaboratory of Molecular Biology and Animal Breeding, School of Animal Science and Technology, Huazhong Agricultural University, Wuhan; ^bDepartment of Gene and Cell Engineering, Institute of Animal Science, Chinese Academy of Agricultural Sciences, Beijing (PR China)

Manuscript received 13 September 2005; accepted in revised form for publication by M. Schmid, 13 October 2005.

Rationale and significance

CRSP9 (cofactor required for Sp1 transcriptional activation, subunit 9) encodes a protein that is a subunit of the CRSP (cofactor required for SP1 activation) complex, which along with TFIID is required for efficient activation by SP1. This protein is also a component of other multisubunit complexes e.g. thyroid hormone receptor (TR) associated proteins that interact with TR and facilitate TR function on DNA templates in conjunction with initiation factors and cofactors (Ryu et al., 1999). The eukaryotic translation termination factor 1 (*ETF1*) gene encodes a class-1 release factor, eRF1, which catalyses termination of protein synthesis at all three stop codons (Inge-Vechtomov et al., 2003). Transmembrane protein 59 (*TMEM59*), also named chromosome 1 open reading frame 8 (*C1orf8*), is a widely expressed protein with function unknown. In this study, the cDNA clones encoding pig *CRSP9*, *ETF1* and *TMEM59* were identified in a cDNA library and we have mapped them to porcine chromosomes using IMpRH panels (Yerle et al., 1998).

Materials and methods

Several porcine cDNA clones were isolated from Chinese Tongcheng pig 55-day fetal longissimus dorsi muscle cDNA library and sequenced (Zhu et al., 2005). After BLAST search in the GenBank non-redundant (nr) sequence database (<http://www.ncbi.nlm.nih.gov/BLAST/Blast.cgi>), we found and identified these three porcine orthologous genes. Primers to amplify across the putative intron regions were designed based on the pig cDNA sequences after alignment of the pig and human sequences. PCR products were confirmed by sequencing and then pig-specific primers were designed in intron regions. Chromosome location was identified by PCR analysis of the porcine whole genome hybrid panel.

Each PCR was performed in duplicate according to the protocol described by Wang et al. (2005). The PCR profile included 3 min at 94°C; 35 cycles of 20 s at 94°C, 20 s at 62°C, and 20 s at 72°C, and a final 5 min extension at 72°C. Data analysis was performed through the IMpRH web site (<http://www.toulouse.inra.fr/lgc/pig/RH/IMpRH.htm>).

Results

We used information from the homologous human gene to deduce the porcine gene structure. The cDNA clones containing the full coding regions of *CRSP9* and *TMEM59* were further analyzed. The porcine *CRSP9* mRNA (GenBank accession number DQ200863) contained a 702-bp ORF flanked by 52 bp 5'UTR and 413 bp 3'UTR and shows high sequence identity (92%) across the CDS (coding sequence) of human *CRSP9*. Computer analysis indicated that porcine *CRSP9* encoded a 233-amino acid protein with a calculated molecular mass of 27.16 kDa and an isoelectric point (pI) of 5.61. The porcine *TMEM59* mRNA (GenBank accession number DQ200860) contained a 972-bp ORF flanked by 119 bp 5'UTR and 476 bp 3'UTR. Porcine *TMEM59* encoded a 323-amino acid protein with a calculated molecular mass of 36.13 kDa and pI of 4.85. The porcine *TMEM59* gene also shows high sequence identity (93%) across the CDS region with that of the human gene and indicates that the gene is conserved

Supported by the Key Project of National Natural Science Foundation of China (30330440), the Key Project of National Basic Research and Developmental Plan of China (G2000016103), National High Science and Technology Foundation of China (2004AA222170) and National Natural Science Foundation of China (30400311). We are grateful to Dr. Martine Yerle for providing the RH panel.

Corresponding author: Kui Li
Department of Gene and Cell Engineering, Institute of Animal Science
Chinese Academy of Agricultural Sciences
Beijing 100094 (PR China)
telephone: +86 10 62833312; fax: +86 10 62899982
e-mail: lkxblghi@public.wh.hb.cn

```

1      CGA AGG CGG TTA CCA GTG TAA ACC CAG TAC CGA GGT TTG TGA TCT 45
46      TCT GCA ATG GGT GAG CCA CAG CAA GTG AGT GCC CTT CCG CCA CCT 90
          M  G  E  P  Q  Q  V  S  A  L  P  P  P 13
91      CCG ATG CAG TAC ATC AAG GAA TAC ACA GAT GAA AAC ATT CAG GAA 135
14      P  M  Q  Y  I  K  E  Y  T  D  E  N  I  Q  E 28
136     GGC CTC GCT CCC AAG CCT CCA CCT CCA ATA AAA GAC AGT TAC ATG 180
29      G  L  A  P  K  P  P  P  P  I  K  D  S  Y  M 43
181     ATG TTT GGC AAC CAG TTC CAA TGT GAT GAT CTT ATC ATC CGC CCT 225
44      M  F  G  N  Q  F  Q  C  D  D  L  I  I  R  P 58
226     TTG GAA AGT CAG GGT ATC GAA CGG CTT CAT CCT ATG CAG TTT GAT 270
59      L  E  S  Q  G  I  E  R  L  H  P  M  Q  F  D 73
271     CAC AAG AAA GAA CTG AGA AAA CTC AAT ATG TCT ATC CTT ATT AAT 315
74      H  K  K  E  L  R  K  L  N  M  S  I  L  I  N 88
316     TTC TTA GAT CTC TTA GAT ATC TTG ATA AGG AGT CCT GGG AGT ATA 360
89      F  L  D  L  L  D  I  L  I  R  S  P  G  S  I 103
361     AAA CGG GAA GAG AAG CTA GAA GAT CTT AAG CTG CTT TTT GTG CAT 405
104     K  R  E  E  K  L  E  D  L  K  L  L  F  V  H 118
406     GTA CAT CAT CTC ATA AAT GAA TAC CGG CCC CAC CAA GCC AGA GAG 450
119     V  H  H  L  I  N  E  Y  R  P  H  Q  A  R  E 133
451     ACC CTG AGA GTC ATG ATG GAG GTG CAG AAG CGG CAA CGC CTT GAG 495
134     T  L  R  V  M  M  E  V  Q  K  R  Q  R  L  E 148
496     ACA GCA GAG AGG TTT CAG AAG CAT CTG GAA CGA GTC ATT GAG ATG 540
149     T  A  E  R  F  Q  K  H  L  E  R  V  I  E  M 163
541     ATT CAG AAT TGC TTG GCT TCT TTG CCT GAT GAC TTG CCG CAT TCA 585
164     I  Q  N  C  L  A  S  L  P  D  D  L  P  H  S 178
586     GTA GCA GGG ATG AGA GTA AAA GCT GAA CCA ATG GAC GCT GAT GAT 630
179     V  A  G  M  R  V  K  A  E  P  M  D  A  D  D 193
631     AGC AAC AAT TGT ACT GGA CAG AGT GAC CAA CAA AGA GAA AAT TCA 675
194     S  N  N  C  T  G  Q  S  D  Q  Q  R  E  N  S 208
676     GGT CAT AGG AGA GAC CAG ATT ATA GAG AAA GAT GCT GCC TTG TGT 720
209     G  H  R  R  D  Q  I  I  E  K  D  A  A  L  C 223
721     GTC CTC ATT GAT GAA ATG AAT GAA AGA CCG TGA AAG ATG TTT CTT 765
224     V  L  I  D  E  M  N  E  R  P  *
766     TTC CTT GCT GAT TTT TTC CTT TCA GTA AAT AGC ATA TAT ATT AGT 810
811     TAA TTC ACT TTT AGA GAA TTC TTT GAA GAG ATA CAA CTA GTA GAA 855
856     GGC ACG CGA ACA GCT TGA CTC TCC AGC TTT ATC AGT TGT GGG ACG 900
901     TCA CAG GCA GTA TTT TCG TCT GTG GCA AGC ATA GCT AGA TTC CTA 945
946     CTT AGC TTA AAT GTA GTG CCC CAG TGA GAG TAA TCA CCG TGA AGG 990
991     GCA GTG TTC GCT GCT ATT GGG TGG AAC GTC CCT GTT TTC AGA GGT 1035
1036    GCA GTG AGC TGG AAG ATA TCC TCA GTA ATT GCA AGC TTT CGT CCT 1080
1081    TGG AAA GTA GAT TTG TCA TTT GCT TTA AGA GTG ATA AAT AAT AAA 1125
1126    GGA TAT CAA ACT ATG TAA ATG TGA AAA AAA AAA AAA AAA 1164

```

Fig. 1. Nucleotide and deduced amino acid sequence of pig *CRSP9*.

Table 1. The chromosome assignments of the three porcine genes

Gene	GenBank Acc. No. (Pig) ^a	Primer sequences (5'-3')	Size (bp)	Human cytogenetic position ^b	Porcine RH mapping results				
					retention (%)	closest marker	dist (cR)	LOD score	porcine cytogenetic position
<i>CRSP9</i>	DQ200861	PL: ATCTTATCATCCGCCCTTTG PR: GTTGCTATCATCAGCGTCCA	428	5q33.3	37	SWR2480	31	12.19	16q22
<i>ETF1</i>	DQ200864	PL: TGCACCTTTTGGCACACTCC PR: AAAGCGCAAGGCTGACTGAC	433	5q31.1	23	IL4	44	8.2	2q28
<i>TMEM59</i>	DQ200862	PL: CACCACAGTTGGAGCAGGAA PR: GTTCCCAACTAGCAGCAACA	336	1p36→p31	35	SW322	21	16.66	6q31→q35

^a The fragments employed for RH mapping.
^b From <http://www.ncbi.nlm.nih.gov/gene/>

3	ACC	GCG	GAG	GCG	GAA	GCA	GCA	GCC	CGT	TGG	GCG	GGG	AAG	GGG	CTG	47
48	GTC	CTG	GAA	GGA	GGA	GGA	AAC	CTT	GAA	AAG	AAA	ACA	GCA	ACA	AGC	92
93	TGA	GCT	GCT	GTG	ACA	GAG	GCA	AAC	AAA	ATG	GCG	GCG	CCG	AAA	GGG	137
										M	A	A	P	K	G	5
138	AGC	CTC	TGG	GTG	CGG	GCC	CAA	CTG	GGG	CTC	ATG	CCG	CTG	CTG	TTG	182
6	S	L	W	V	R	A	Q	L	G	L	L	P	L	L	L	20
183	CTG	ACC	ATG	GCC	CTG	GCC	GGA	GGC	CCA	GGG	ACT	GCT	TCC	GCT	GAA	227
21	L	T	M	A	L	A	G	G	P	G	T	A	S	A	E	35
228	GCG	TTT	GAC	TCG	GTG	TTG	GGT	GAT	ACG	GCA	TCT	TGT	CAC	GCG	GCC	272
36	A	F	D	S	V	L	G	D	T	A	S	C	H	R	A	50
273	TGT	CAG	TTG	ACC	TAC	CCC	TTG	CAC	ACC	TAC	CCC	AAG	GAG	GAG	GAG	317
51	C	Q	L	T	Y	P	L	H	T	Y	P	K	E	E	E	65
318	TTG	TAC	GCC	TGT	CAG	AGA	GGT	TGC	AGG	CTG	TTT	TCC	ATT	TGT	CAG	362
66	L	Y	A	C	Q	R	G	C	R	L	F	S	I	C	Q	80
363	TTT	GTG	GAT	GAT	GGA	ATT	GAC	TTA	AAT	CGG	ACC	AAA	CTG	GAA	TGT	407
81	F	V	D	D	G	I	D	L	N	R	T	K	L	E	C	95
408	GAA	TCT	GCA	TGT	ACA	GAA	GCA	TAT	TCC	CAA	TCT	GAT	GAG	CAA	TAT	452
96	E	S	A	C	T	E	A	Y	S	Q	S	D	E	Q	Y	110
453	GCC	TGC	CAT	CTT	GGT	TGC	CAG	AAT	CAG	CTG	CCA	TTT	GCT	GAG	CTG	497
111	A	C	H	L	G	C	Q	N	Q	L	P	F	A	E	L	125
498	AGA	CAA	GAG	CAA	CTC	ATG	TCC	CTG	ATG	CCA	AAA	ATG	CAT	CTA	CTC	542
126	R	Q	E	Q	L	M	S	L	M	P	K	M	H	L	L	140
543	TTC	CCT	TTA	ACT	CTG	GTG	AGG	TCA	TTC	TGG	AGT	GAC	GTG	ATG	GAC	587
141	F	P	L	T	L	V	R	S	F	W	S	D	V	M	D	155
588	TCG	GCA	CAG	AGC	TTT	ATA	ACC	TCT	TCA	TGG	ACT	TTT	TAT	CTT	CAA	632
156	S	A	Q	S	F	I	T	S	S	W	T	F	Y	L	Q	170
633	GCT	GAT	GAT	GGG	AAA	ATA	GTT	ATC	TTC	CAG	TCC	AAG	CCA	GAA	ATC	677
171	A	D	D	G	K	I	V	I	F	Q	S	K	P	E	I	185
678	CAG	TAT	GCA	CCA	CAG	TTG	GAG	CAG	GAA	CCT	ACA	AAT	TTG	AAA	GAA	722
186	Q	Y	A	P	Q	L	E	Q	E	P	T	N	L	K	E	200
723	TCG	TCA	CTG	AGC	AAA	ATG	TCC	TAT	TTG	CAA	ATG	AGA	AGT	TCA	CAA	767
201	S	S	L	S	K	M	S	Y	L	Q	M	R	S	S	Q	215
768	GCA	CAC	AGG	AAC	TAT	CTT	GAA	GAT	GGG	GAA	AAC	GAT	GGC	TTT	TTA	812
216	A	H	R	N	Y	L	E	D	G	E	N	D	G	F	L	230
813	AGA	TGC	CTC	TCT	CTT	AAC	TCT	GGG	TGG	ATT	TTA	ACT	ATG	ACT	CTT	857
231	R	C	L	S	L	N	S	G	W	I	L	T	M	T	L	245
858	GTC	CTC	TCC	GTG	ATG	GTA	TTG	CTC	TGG	ATT	TGT	TGT	GCA	ACT	GTT	902
246	V	L	S	V	M	V	L	L	W	I	C	C	A	T	V	260
903	GCT	ACA	GCT	GTG	GAG	CAG	TAT	GTT	CCC	TCT	GAG	AAG	CTG	AGT	ATC	947
261	A	T	A	V	E	Q	Y	V	P	S	E	K	L	S	I	275
948	TAT	GGT	GAC	TTG	GAA	TTT	GTG	AAT	GAA	CAA	AAG	CTA	AAC	AGA	TAT	992
276	Y	G	D	L	E	F	V	N	E	Q	K	L	N	R	Y	290
993	CCA	GCT	TCT	TCT	CTT	GTG	CTT	GTT	AGA	TCT	AAA	GCT	GAA	GAT	CAT	1037
291	P	A	S	S	L	V	L	V	R	S	K	A	E	D	H	305
1038	GAC	GAA	GCA	GGG	CCA	TTA	CCT	ACA	AAA	GTG	AAT	CTT	GCT	CAT	TCA	1082
306	D	E	A	G	P	L	P	T	K	V	N	L	A	H	S	320
1083	GAA	ATT	TAA	GCT	TTT	TTT	TTT	TTT	TAA	AGA	AAA	AAT	GTA	ATA	GAC	1127
321	E	I	*													
1128	ATC	TAA	ATT	TCC	ATT	CCT	CAT	AGA	GCT	TTT	TAA	GAT	GGT	TTT	GTT	1172
1173	GGA	TAT	AGG	GCT	TAA	AAA	TCA	CTA	TTA	AAT	GCA	AAT	AAA	GTT	ACC	1217
1218	CAA	ATC	TGT	GAA	GAC	TGT	ATT	TGC	TGT	AAC	TTT	ACC	TAT	AAT	TTT	1262
1263	TTT	TCT	AGT	AAT	TTA	AGA	GAA	GGA	TAT	TTG	GGA	TTG	TAT	TTT	TAT	1307
1308	TTT	ACT	AAT	CTC	TGT	AGC	TAT	TGT	AAT	TGT	TAT	TTG	CAT	TGT	TTT	1352
1353	TGT	TTG	TTT	CTT	TCT	TAG	TGA	AAT	TCT	AGG	AGC	TGA	GCA	TTG	TCC	1397
1398	TTC	CCC	ACC	TCC	TGT	GGT	GAT	ACT	GTG	AGT	CAC	CAG	CTG	AAT	ACT	1442
1443	CAG	TAG	GAT	ATG	ATG	CTC	CTA	CTC	AGA	AAT	GGC	CCA	CAA	CCT	GTA	1487
1488	ATT	TGA	AAT	GTA	GCA	GAA	AAT	GCC	CTT	TAA	TGA	CAC	TAC	ATT	TTG	1532
1533	AGT	TGT	ATG	GAA	CTG	AGA	TCA	TTA	AAA	TTT	TGT					1565

Fig. 2. Nucleotide and deduced amino acid sequence of pig *TMEM59*.

through evolution. The molecular characterization of porcine *TMEM59* may provide a clue for functional studies of the human *TMEM59* gene.

As shown in Table 1, the pig *CRSP9*, *ETF1* and *TMEM59* genes were assigned to porcine chromosomes 16, 2 and 6 under the minimum threshold of 5 (LOD score), respectively. The cytogenetic locations were deduced from the closest markers available in the cytogenetic map (Hawken et al., 1999).

According to bi-directional painting of human and porcine chromosomes (Goureau et al., 1996), HSA5q14→q35 was expected to correspond to SSC2. *CRSP9* and *ETF1* are very close together in both human and mouse, however, they have been mapped to different porcine chromosomes (SSC16 and SSC2) in this experiment. Another gene *GRIAI* (glutamate receptor, ionotropic, AMPA 1) which is located in the 5q33 region in human has been assigned to SSC16 (q14 or q22→q23) in pig and suggests that this homologous region was in a rather small gap so the signal could not be detected by painting (Larsen et al., 1999). The assignment of porcine *CRSP9* may also prove this.

References

- Goureau A, Yerle M, Schmitz A, Riquet J, Milan D, Pinton P, Frelat G, Gellin J: Human and porcine correspondence of chromosome segments using bi-directional chromosome painting. *Genomics* 36:252–262 (1996).
- Hawken RJ, Murtaugh J, Flickinger GH, Yerle M, Robic A, Milan D, Gellin J, Beattie CW, Schook LB, Alexander LJ: A first-generation porcine whole genome radiation hybrid map. *Mamm Genome* 10:824–830 (1999).
- Inge-Vechtomov S, Zhouravleva G, Philippe M: Eukaryotic release factors (eRFs) history. *Biol Cell* 95:195–209 (2003).
- Larsen NJ, Marklund S, Kelly KA, Malek M, Tuggle CK, Yerle M, Rothchild MF: New insights into porcine-human synteny conservation. *Mamm Genome* 10:488–491 (1999).
- Ryu S, Zhou S, Ladurner AG, Tjian R: The transcriptional cofactor complex CRSP is required for activity of the enhancer-binding protein Sp1. *Nature* 397:446–450 (1999).
- Wang H, Zhu ZM, Yerle M, Wu X, Wang HL, Gu MS, Li K: Assignment of three novel genes to porcine chromosome 13 by a radiation hybrid panel. *Cytogenet Genome Res* 108:363J (2005).
- Yerle M, Pinton P, Robic A, Alfonso A, Palvadeau Y, Delcros C, Hawken R, Alexander L, Beattie C, Schook L, Milan D, Gellin J: Construction of a whole-genome radiation hybrid panel for high-resolution gene mapping in pigs. *Cytogenet Cell Genet* 82:182–188 (1998).
- Zhu ZM, Zhang JB, Li K, Zhao SH: Cloning, mapping and association study with carcass traits of the porcine *SDHD* gene. *Anim Genet* 36:191–195 (2005).